DOCUMENT RESUME

ED 422 390 TM 028 952

AUTHOR Thompson, Bruce; Fan, Xitao
TITLE Bootstrap Estimation of Sample Statistic Bias in Structural Equation Modeling.
PUB DATE 1998-04-00
NOTE 40p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Estimation (Mathematics); Goodness of Fit; Monte Carlo Methods; Sample Size; *Statistical Bias; *Structural Equation Models
IDENTIFIERS *Bootstrap Methods

ABSTRACT

This study empirically investigated bootstrap bias estimation in the area of structural equation modeling (SEM). Three correctly specified SEM models were used under four different sample size conditions. Monte Carlo experiments were carried out to generate the criteria against which bootstrap bias estimation should be judged. For SEM fit indices, bias estimates from the bootstrap and Monte Carlo experiments were quite comparable in most cases. It is noted that bias was constrained in one direction in the Monte Carlo experiments because of the perfect fit of the true SEM models. For the SEM loadings and coefficients, the difference between bootstrap and Monte Carlo bias estimations was very small, and the distributions of the bias estimators from the two experiments were quite similar. For the SEM variances/covariances, the comparison of the bias estimator distributions from the two experiments indicated that bootstrap bias estimation could be considered adequate. Because the study involved three SEM models which served as an internal replication mechanism, the likelihood of chance discovery for the findings was small, and the findings should have reasonable generalizability. Future studies may extend the current findings by examining misspecified SEM models. Data nonnormality may be another dimension to be considered in future investigations. (Contains 6 figures, 4 tables, and 40 references.) (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

BTRAPSEM.DOC

ED 422 390

BOOTSTRAP ESTIMATION OF SAMPLE STATISTIC BIAS IN STRUCTURAL EQUATION MODELING

Xitao Fan

Utah State University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Xitao Fan
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Bruce Thompson

Texas A&M University
and
Baylor College of Medicine

Running Head: Bootstrap Bias Estimation in SEM

Note: Correspondence about this paper should be sent to the first author at: Dept. of Psychology, Utah State University, Logan, UT 84322-2810. Phone: (435)797-1451, and E-Mail: fafan@cc.usu.edu

TM028952

Paper presented at the 1998 Annual Meeting of the American Educational Research Association, San Diego, CA, April 13-17 (*Session #: 15.51*).

ABSTRACT

This study empirically investigated bootstrap bias estimation in the area of structural equation modeling. Three correctly specified SEM models were used under four different sample size conditions. Monte Carlo experiments were carried out to generate the criteria against which bootstrap bias estimation could be judged. For SEM fit indices, bias estimates from the bootstrap and Monte Carlo experiments were quite comparable for most of them It is noted that bias was constrained in one direction in the Monte Carlo experiments because of the perfect fit of the true SEM models. For the SEM loadings and coefficients, the difference between bootstrap and Monte Carlo bias estimations was very small, and the distributions of the bias estimators from the two experiments were quite similar. For the SEM variances/covariances, the comparison of the bias estimator distributions from the two experiments indicated that bootstrap bias estimation could be considered adequate. Because the study involved three SEM models which served as internal replication mechanism, the likelihood of chance discovery for the findings was small, and the findings should have reasonable generalizability. Future studies may extend the current findings by examining misspecified SEM models. Data non-normality may be another dimension to be considered in future investigation.

Structural equation modeling (SEM) has increasingly been seen as a useful quantitative technique for specifying, estimating, and testing hypothesized models describing relationships among a set of substantively meaningful variables. Much of SEM's attractiveness is due to the method's applicability in a wide variety of research situations, a versatility that has been amply demonstrated (e.g., Bollen & Long, 1993; Byrne, 1994; Jöreskog & Sörbom, 1989; Loehlin, 1992; SAS Institute, 1990).

Furthermore, many widely used statistical techniques may also be considered as special cases of SEM, including regression analysis, canonical correlation analysis, confirmatory factor analysis, and path analysis (Bagozzi, Fornell & Larcker, 1981; Bentler, 1992; Fan, 1996; Jöreskog & Sorbom, 1989). Because of such generality, SEM has been heralded as a unified model which joins methods from econometrics, psychometrics, sociometrics, and multivariate statistics (Bentler, 1994). In short, for researchers in the social and behavioral sciences, SEM has become an important tool for testing theories with both experimental and non-experimental data (Bentler & Dudgeon, 1996).

Bootstrap method, on the other hand, has also been applauded as one of the newest breakthroughs in statistics (Kots & Johnson, 1992). The significance of bootstrapping as a versatile nonparametric statistical approach to data analysis has been widely recognized not only by those in the area of statistics, but also by quantitative researchers in social and behavioral sciences. In the area of education, the recognition for the bootstrapping method was evidenced by the invited keynote address delivered by the pioneer in bootstrapping, Bradley Efron, at the 1995 AERA Annual Meeting (Efron, 1995).

In social and behavioral sciences, bootstrap method has been used in a variety of research situations, and for many different statistical techniques. For example, bootstrap method has been

applied in psychological measurement for such issues as differential test predictive validity (e.g., Fan & Mathews, 1994) and item bias (e.g., Harris & Kolen, 1991), and in sociological research (e.g., Stine, 1989). The application of bootstrapping method has involved many different statistical techniques, including correlation analysis (e.g., Rasmussen, 1987), regression analysis (e.g., Fan & Jacoby, 1995), discriminant analysis (e.g., Dalgleish, 1994), canonical correlation analysis (e.g., Fan & Wang, 1996), factor analysis (e.g., Lambert, Wildt, & Durand, 199; Thompson, 1988) , and structural equation modeling (e.g., Bollen & Stine, 1990). In addition to its original use as a nonparametric alternative to statistical significance testing (Efron, 1985), bootstrap technique has also been advocated as a method of internal replication to assess the replicability of results of an individual study (Thompson, 1993).

As a general nonparametric technique, the application of bootstrap may be most appropriate in situations where the statistical theory is weak, or where theoretical assumptions are unlikely to be tenable. Structural equation modeling can be considered as an area in which the statistical theory is relatively weak (e.g., no theoretical sampling distributions for many model fit indices), and in which theoretical assumptions are often untenable (e.g., multivariate normality is often violated). Viewed inn this perspective, the application of bootstrap method in SEM can be considered quite relevant and desirable.

It has been observed that the application of bootstrap in SEM could be categorized for four different purposes: (1) to estimate bias of sample statistic (bias estimation), (2) to estimate standard errors, (3) to construct confidence intervals, and (4) to test SEM models (Yung & Bentler, 1996). Despite its conceptual and procedural simplicity, and its versatility in a variety of statistical situations, the success of applying bootstrap method as originally proposed (Efron, 1979) in SEM may not be guaranteed. For example, it has been shown that bootstrap application

(as originally defined) has failed in the area of SEM model testing (Bollen & Stine, 1993). Because of the uncertainty about its success in SEM, Yung and Bentler (1996) made a general suggestion that, in the area of SEM, bootstrap results should not be blindly and indiscriminantly trusted; instead, a particular application of bootstrapping (e.g., bias estimation) in SEM needs to be investigated until the validity of such application has been reasonably established.

Of the four major applications of bootstrap in SEM, bias estimation (i.e., to use bootstrap-estimated bias of a statistic to represent the true bias of that statistic) has been identified as an area where little empirical evidence exists (Yung & Bentler, 1996). Consequently, the validity of bootstrap-based bias estimation in SEM is largely unknown. The problem here is that no studies have been reported which investigated bootstrap bias estimation for SEM. The most relevant studies which investigated bootstrap-based bias estimation actually applied bootstrap in exploratory factor analysis models, but not SEM models per se.

Chatterjee (1984), based on the empirical results from exploratory factor analysis, concluded that the bootstrap-estimated bias was very small for loadings in exploratory factor analysis. The results of the study, however, may not be very meaningful, for the reason that there was no external criterion against which the validity of the bootstrap-based bias estimation could be judged; instead, the researcher relied entirely on his "faith" in the bootstrap principle (Yung & Bentler, 1996). The study by Ichikawa and Konishi (1995) provided the needed external criteria ("true" bias estimation empirically obtained through Monte Carlo simulation) for judging the adequacy of bootstrap-based bias estimation. Their results indicated that bootstrap-based bias estimation worked well for rotated factors, but less well for unrotated factors. Although this study represents substantial improvement over that of Chatterjee (1984), the fact that both studies investigated exploratory factor analysis, not confirmatory factor analysis or full SEM models,

considerably limited the generalizability of their results for SEM. As discussed by Yung and Bentler (1996), ". . . bootstrap estimator for biases may not work equally well in all cases .... It would seem risky indeed to have blind faith in the bootstrap principle in the hope that bootstrap estimation of biases in the context of structural equation modeling will work out correctly." (p. 201).

In the most general form, the bias of an estimator $\hat{\theta}$ for θ is expressed as

$$B(\hat{\theta}) = \varepsilon(\hat{\theta}_{\Delta}) \tag{1}$$

where θ is the population parameter of any sort, $\hat{\theta}$ is the sample statistic for θ, and $\hat{\theta}_{\Delta}$ is the deviation of the sample statistic from the population parameter θ (i.e., $\hat{\theta}_{\Delta} = \hat{\theta} - \theta$), and $\varepsilon(\hat{\theta}_{\Delta})$ is the expected value of $\hat{\theta}_{\Delta}$. In an application, the true bias of the statistic $B(\hat{\theta})$ is usually estimated through Monte Carlo simulation. When N samples are drawn from the specified population, we have each individual sample statistic deviation $\hat{\theta}_{\Delta j}$:

$$\hat{\theta}_{\Delta j} = \hat{\theta}_{j} - \theta \tag{2}$$

The bias of the statistic is estimated through:

$$\begin{aligned} B(\hat{\theta}) &= \bar{\theta}_{\Delta} \\ &= \frac{1}{N} \sum \hat{\theta}_{\Delta j} \quad (j=1, 2, 3, \ldots N) \end{aligned} \tag{3}$$

In other words, bias estimation is the averaged value of $\hat{\theta}_{\Delta j}$ obtained from repeated Monte Carlo sampling from the defined population.

In bootstrap, the bias of a bootstrapped sample statistic is defined as the difference ($\hat{\theta}_{\Delta j}^{*}$) between the parent sample statistic ($\hat{\theta}$) and that obtained from a bootstrapped sample ($\hat{\theta}_{j}^{*}$) .(i.e., $\hat{\theta}_{\Delta j}^{*} = \hat{\theta}_{j}^{*} - \theta$). When B samples are bootstrapped from the parent sample through sampling with replacement, the bootstrap bias estimation is obtained by:

$$B^{*}(\hat{\theta}) = \bar{\theta}_{\Delta}^{*} = \frac{1}{B} \sum \hat{\theta}_{\Delta j}^{*} \quad (j=1, 2, 3, \ldots B) \tag{4}$$

Obviously, bootstrap principle implies that $\bar{\theta}_{\Delta}$ in (3) is comparable to the $\bar{\theta}_{\Delta}^{*}$ in (4).

In addition to the issue of whether or not bootstrap-based bias estimation (Equation 4) is systematically comparable with the estimated true bias of a statistic as defined in Equation 3, another relevant issue is about the distribution of bootstrap-based bias estimator. It is possible that bootstrap-based bias estimation may be comparable with the true bias of a statistic, but the variability of the bootstrap-based bias estimator ($\hat{\theta}_{\Delta j}^{*}$) is substantially different from that of the bias estimator ($\hat{\theta}_{\Delta j}$) from Monte Carlo experiments. In other words, comparability between the bootstrap-based bias estimation and the Monte Carlo bias estimation does not necessarily mean the comparability of distributions of the bias estimators from the two situations.

The present study was designed to provide empirical information with regard to the adequacy of bootstrap-based bias estimation in structural equation modeling. The study was concerned with two research questions:

(1) How comparable is the bootstrap-based bias estimation with the "true" bias (as estimated from Monte Carlo simulation)?

(2) How comparable is the distribution of the bootstrap bias estimator with the distribution of

the bias estimator from Monte Carlo experiments?

Both these two research questions were investigated with regard to:

(a) SEM fit indices,

(b) factor loadings and path coefficients in SEM models,

(c) variances and covariances (for both observed and latent variables) in SEM models.

METHODS

Several issues were considered in the design of this study: the generalizability of the results of the study, the issue of sample size, and the criterion for judging the adequacy of the bootstrap bias estimation. Attempts were made to accommodate these relevant issues in the study design.

Increasing the Generalizability of Findings

Two features were incorporated in the study design for the purpose of increasing the generalizability of the study results: multiple SEM models, and realistic SEM models and model parameters. Because it is always uncertain to what extent the results from a particular SEM model would capture the complexity of SEM modeling in general, it is possible that the findings based on a particular model may reflect some idiosyncracies associated with the model, and consequently, may only have limited generalizability. In this study, to minimize the possibility that the idiosyncracies of a particular model may be mistaken for something more generalizable, three different SEM models (one confirmatory factor analysis model and two full structural models) with varying degrees of model complexity were simulated. Such a design of multiple models provided a very useful mechanism for verifying the internal validity of the findings. As is generally recognized, internal validity is the necessary, although not sufficient condition of external validity

(i.e., the generalizability of the findings) (Gall, Borg, & Gall, 1996).

In addition to the multiple SEM models, this study adopted realistic SEM models which were based on substantive research studies. As suggested by Gerbing and Anderson (1993), simulating substantively meaningful models might increase the external validity of a study. Not only were all the three SEM models simulated in this study from substantive research studies, but the model parameters simulated also closely matched those in the substantive studies. Of the three SEM models simulated in this study, one is a confirmatory factor analysis model originally appeared in the research report by Calsyn and Kenny (1977), and later discussed as a confirmatory factor analysis model example by Jöreskog and Sörbom (1989, p. 83). The second SEM model is a full structural equation model discussed by Jöreskog and Sörbom (1989, p. 178) which was based on some longitudinal data from a study conducted by the Educational Testing Service. The third model is also a full structural equation model which originally appeared in the work by Wheaton, Muthén, Alwin, and Summers (1977), and which was later widely discussed in the SEM literature (e.g., Bentler, 1992; Jöreskog & Sörbom, 1989; SAS Institute, 1990). To limit the scope of the study, only statistically true models are simulated, although, ideally, models with model specification error should also be studied in the future. Figure 1, Figure 2, and Figure 3 present the three SEM models used in this study, and their respective model parameters in LISREL terms.

Insert Figure 1, 2, 3 about here

Sample Size

It is unclear how large a sample should be in SEM applications. The research findings on

this issue are inconclusive (MacCallum, Roznowski, & Necowitz, 1992; Tanaka, 1987). It has been reported that small sample size led to not only untrustworthy estimation results and fit indices, but also the high rate of improper solutions in simulation (Ichikawa & Konishi, 1995). Sample size of 200 in SEM applications has been considered as being relatively small by many (Boomsma, 1982; Camstra & Boomsma, 1992; Ichikawa & Konishi, 1995; MacCallum, et al., 1992). Some researchers even consider sample sizes in the thousands to be adequate (e.g., Hu, Bentler, & Kano, 1992; Marsh et al., 1988). Realistically, however, such large sample sizes are often beyond the reach of most researchers. Also, as pointed out by many researchers, it may not be realistic to suggest a single value to define a small or large sample, because models and the number of free parameters vary from application to application, so consideration of sample size should be related to model complexity and the number of free parameters (MacCallum et al., 1992; Tanaka, 1987).

In this study, four sample size conditions (n=100, 200, 500, 1000) were simulated. As discussed above, there are no objective criteria on this issue. Based on our experience and our perception about the current practice in this area, we felt that these sample size conditions were quite representative of the current practice, and the four sample size conditions were also considered to be reasonably spaced.

"True" Bias as the Criteria for Judging Bootstrap Bias Estimation

As discussed in the literature review section, in order to understand the adequacy of bootstrap bias estimation, external criteria must first be obtained against which bootstrap bias estimation can be compared. In this study, the external criteria ("true" bias) were obtained empirically through Monte Carlo simulation. The three SEM models, the four sample size conditions, and 200 replications within each cell made up a total of 2400 (3×4×200 = 2400)

random samples to be fitted to one of the three SEM models. Based on the Monte Carlo experiments, "true" bias for an SEM statistic ($\hat{\theta}$) was empirically obtained through Equations 2 and 3.

Within each cell condition described above, one sample was generated to served as the parent sample for bootstrapping, and 200 samples were bootstrapped from this parent sample using "completely nonparametric bootstrap" approach (Yung & Bentler, 1996). In this bootstrap phase of the study, a total of 2400 samples were obtained through bootstrapping resampling, and fitted to one of the three SEM models. The bootstrap bias estimation for the statistic ($\hat{\theta}$) was obtained through Equation 4. The Monte Carlo bias estimate $\mathbf{B}(\hat{\theta})$ and the bootstrap bias estimate $\mathbf{B}^*(\hat{\theta})$ were empirically compared to check if the bootstrap bias estimation $\mathbf{B}^*(\theta)$ was systematically different than the Monte Carlo bias estimate $\mathbf{B}(\hat{\theta})$. In addition, the sampling distribution of the bootstrap bias estimator ($\hat{\theta}_{\Delta j}^*$) was compared with that of the Monte Carlo bias estimator ($\hat{\theta}_{\Delta j}$) to check how comparable the bias estimator distributions were.

Data generation was accomplished by using the SAS normal data generator. Multivariate normal data were simulated using the matrix decomposition procedure (Kaiser & Dickman, 1962). All data generation, bootstrap resampling, model fitting (SAS PROC CALIS) were accomplished through the SAS system (SAS Version 6.11).

RESULTS AND DISCUSSIONS

Because the amount of results which could be presented was huge, our presentation of the results was selective, and we tried to focus on the main questions. Because all the simulation results have been saved, some additional analyses can be conducted if there is enough interest. Our presentation of the results and the related discussion followed the order of the research

questions presented earlier in the paper.

Comparability between Bootstrap Bias Estimation and Monte Carlo Bias Estimation

The first research question asks, "How comparable is the bootstrap bias estimation with the "true" bias in SEM?" To answer this question, we presented the results separately for (1) SEM fit indices; (2) loadings and path coefficients in SEM; and (3) variances and covariances in SEM (for both observed and latent variables).

SEM fit indices. Table 1 presents the bootstrap bias estimation [$\mathbf{B}^*(\hat{\theta})$] and the Monte Carlo bias estimation [$\mathbf{B}(\hat{\theta})$] for a variety of SEM fit indices. A negative sign indicates a downward bias (the estimate smaller than the parameter), and a positive sign (no sign) indicates an upward bias (the estimate larger than the parameter). A close look at the table reveals several observations. First, both the bootstrap and the Monte Carlo has downward bias for all the fit indices except the χ^2 and the RMR (root mean square residual), for which there is upward bias. For Monte Carlo results, this is expected, because all the models are true models, and the fit for the defined population is perfect. In other words, for true SEM models, the population parameter values of these indices function either as the "ceiling" (all except χ^2 and RMR) or the "floor" (χ^2 and RMR) of the possible values of the statistics. As such, bias of the statistic is constrained to be in one direction only: downward (for fit indices with a "ceiling") or upward (for indices with a "floor").

The fact that the bootstrap bias estimation is consistently in the same direction as the "true" bias (from the Monte Carlo results) for all the fit indices is somewhat remarkable. This is because, as Bollen and Stine (1993) discussed, the bootstrap resampling space (i.e., the bootstrap parent sample) does not represent the null model (i.e., the correctly specified true model). This fact means that, for these fit indices, the bootstrap parent sample statistic may not function as the

"ceiling" or "floor", as the population parameters do in Monte Carlo simulation; and some bootstrapped samples may have better fit index values than those of the parent sample. In spite of the difference in this regard, bootstrap still produces bias estimation consistent in direction with the Monte Carlo estimated bias.

The second observation from Table 1 is that, for most of the fit indices, bootstrap bias estimation is also quite comparable in magnitude with the Monte Carlo estimated bias, probably with the exception of RMR, CFI, CENTRA, NNFI, and DELTA2. The comparability in magnitude for χ^2, GFI, AGFI, NFI, RHO1, PGFI, and PARMS are quite remarkable. The other five indices, however, showed less comparability between the two kinds of experiments. Of these five fit indices, bootstrap estimated bias is consistently smaller in magnitude than the bias from the Monte Carlo experiments for RMR, but bootstrap estimated bias tends to be larger than the Monte Carlo bias for the other four.

In general, the bootstrap method performed quite well in bias estimation for most of the SEM fit indices examined above when compared with the bias estimated from Monte Carlo experiments. As expected, the magnitude of bias decreases with the increase of sample size for both the bootstrap estimated bias and the "true" bias. It is important to note that the findings discussed above are evident from the three SEM models with different data characteristics, so it is unlikely that these observations may be the results of the idiosyncracies of any particular model or the data structure associated with it.

Loadings and path coefficients in SEM. Table 2 presents the comparison of bootstrap bias estimation with the "true" bias for the factor loadings and coefficients of the three SEM models (Λ_X, Λ_Y, Γ, and B matrices in LISREL terms). Because the three models have different loadings and coefficients, Table 2 actually consists of three sub-tables, each for one SEM model.

It should be noted that these are bias estimates for the unstandardized coefficients and loadings, not for the standardized coefficients and loadings. The specified values of these original loadings and coefficients, however, range approximately from 0.2 to 1.2, similar to standardized coefficients in terms of the scale range. Several observations can be made about Table 2. First, the absolute magnitude of the estimated bias (both from the bootstrap and the Monte Carlo experiments) for these statistics is very small for the overwhelming majority of them. Typically, the bias only appears at the third or fourth decimal place for these loadings and coefficients. Only a few entries show bias at the second decimal place, and one or two entries show bias at the first decimal place. The small bias in the majority of the cases suggests that, for practical research applications, bias of these SEM statistics is minimal, and may safely be ignored without much loss of information.

Relatively speaking, however, contrary to the situation about SEM fit indices where high degree of comparability between the bootstrap and Monte Carlo bias estimates was observed for most of the fit indices, there does not appear to be a consistent pattern between the bootstrap and Monte Carlo bias estimates for these loadings and coefficients, either in terms of bias direction, or in terms of bias magnitude. In some cases, bootstrap bias is larger than the Monte Carlo bias, while the reverse is true in others. The same can also be said about the direction of bias in the two situations. It may be argued that the minuscule amount of bias in both situations may have made the relative difference between the two (bootstrap and Monte Carlo bias estimates) unimportant for practical purposes. Again, as expected, the magnitude of bias decreases with the increase of the sample size.

Variances and covariances in SEM. Table 3 presents the bias estimation for the variances and covariances in SEM (θ_δ, θ_ε, Φ, and Ψ matrices in LISREL terms). Like Table 2, Table 3

actually consists of three sub-tables, each for one SEM model. It should be noted that these are estimated biases for the variances and covariances in SEM models, and the specified parameter values of these variances and covariances may be quite large (ranging from 1.6 to 115; see Figures 1, 2, and 3 for model specifications). Because the original values may be quite large, we need to be careful in interpreting the magnitude of these bias estimates. For example, for the coefficient ϕ_{11} in SEM Model 1, the bootstrap estimated bias and the Monte Carlo estimated bias for sample size condition N=100 is approximately 2.5 (though in different directions). This value of bias may sound large if we have not switched from the measurement scales in Table 2. But the specified value for ϕ_{11} was 105, and percentagewise, the bias is only about 2.3% of the parameter value, and comparable to a bias of 0.012 for a correlation coefficient of 0.5.

Like Table 2, there is a lack of consistent pattern between bias estimates from bootstrap and Monte Carlo experiments, both in terms of bias direction (upward or downward) and in terms of bias magnitude. Although, in many cases, bootstrap estimated bias appears to be larger than the Monte Carlo estimated bias in terms of magnitude, the opposite is also abundant in the same table. The increase of the sample size, as in Table 1 and Table 2, has the expected tendency of reducing the magnitude of bias for both bootstrap and Monte Carlo estimated bias.

Another observation about the bootstrap and Monte Carlo experiments is related to the issue of improper solutions. In this study, samples (either in bootstrap or Monte Carlo experiments) with improper solutions were excluded from further analyses, and they were not replaced with new samples. It is observed that the occurrence of improper solutions is higher for bootstrap experiments than for Monte Carlo experiments. The occurrences of improper solutions under different sample size conditions and for the different SEM models are presented in Table 4. It is seen that improper solutions mainly occurred for N=100 and N=200 conditions, especially for

SEM Model 2.

Insert Table 4 about here

Comparability of Bias Estimator Distributions: Bootstrap ($\hat{\theta}_{\Delta j}^{*}$) vs. Monte Carlo ($\hat{\theta}_{\Delta j}$)

The second research question asks, "how comparable is the distribution of the bootstrap bias estimator and that of the Monte Carlo bias estimator?" Because so many different kinds of statistics are involved for different models and under different sample size conditions, it is difficult to describe the distributional characteristics of all the estimators in a concise and easy-to-understand manner. Instead of trying to be complete, we graphically presented a limited number of cases to illustrate our key observations with regard to the distributional characteristics of the bootstrap bias estimators vs. the Monte Carlo bias estimators.

SEM fit indices. Figure 4 presents the comparisons of bootstrap bias estimator distributions with those of Monte Carlo bias estimators for four SEM fit indices [GFI (SEM Model 2, N=100), NFI (CFA Model, N=200), χ^2 (CFA Model, N=200), and RMR (SEM Model 1, N=100)]. In this and in the following Figures 5 and 6, black bars represent the distributions of the bootstrap bias estimator, while the shaded bars represent the distributions of Monte Carlo bias estimator. As discussed previously, because the true SEM models were simulated in the Monte Carlo simulation, these fit indices have either a "ceiling" (GFI and NFI) or a "floor" (χ^2, and RMR). As a result, the bias from the Monte Carlo experiments could only be either all negative (downward bias, for GFI and NFI) or all positive (upward bias, for χ^2, and RMR). On the other hand, for the bootstrap resampling space (i.e., bootstrap parent sample), the models were not true SEM models (see Bollen & Stine, 1993, for the discussion related to the issue). As a result, the

bootstrap bias theoretically could be either positive or negative for all these fit indices. This difference underlies the disparity between the distributions of the bootstrap bias estimator and the Monte Carlo bias estimator. The distributional difference of the bootstrap and Monte Carlo bias estimators is very similar for GFI, NFI, and χ^2, although χ^2 is opposite in direction.

Insert Figure 4 about here

In Table 1 presented earlier, bootstrap bias estimation for RMR is consistently smaller than the "true" bias. Figure 4 (d) reveals the distributional differences of the bootstrap bias estimator and the Monte Carlo bias estimator. Contrary to Figure 4 (a), (b), and (c) where the overlap of the two bias estimator distributions was substantial, for RMR, the two distributions are quite different in terms of both the location and the shape of the distributions, with the location of the bootstrap bias estimator being much closer to zero. This explains why bootstrap bias estimation was much smaller than that based on Monte Carlo simulation as presented in Table 1.

The disparity between the bootstrap bias estimator distribution and that of the Monte Carlo bias estimator, however, may partly be the consequence of simulating true SEM models with their inherent "floor" or "ceiling" for the fit indices, which constrains the bias to be only in one direction for Monte Carlo experiments. But the bootstrap bias estimation is less constrained in this regard, because the bootstrap resampling space represents a misspecified SEM model, rather than a true SEM model as in the Monte Carlo simulation. It is likely that the disparity between the bootstrap bias estimator distribution and that of the Monte Carlo bias estimator would be smaller than what has been observed in Figure 4 if misspecified SEM models are used in the study. This possibility, of course, needs to be empirically examined.

SEM loadings and coefficients. Figure 5 presents the distributional comparisons between

the bootstrap bias estimator and the Monte Carlo bias estimator for four SEM loadings and coefficients [λ_{X21} (CFA Model, N=200), λ_{Y32} (SEM Model 1, N=200), γ_{11} (SEM Model 1, N=500), and β (SEM Model 2, N=100)]. It is seen that, for all the four coefficients, the distributions of the bootstrap bias estimator and the Monte Carlo bias estimator are quite comparable in terms of their range and shape, indicating that the bootstrap bias estimator is close to the Monte Carlo bias estimator. Notice that, for these loadings and coefficients, there do not exist any constraints caused by a theoretical ceiling or floor, as in the case for the fit indices. As a result, bias estimator can be distributed in both directions (positive and negative) for both the bootstrap and the Monte Carlo experiments. The similar distributions of the bootstrap and the Monte Carlo bias estimators support the earlier discussion that the minuscule difference between the two bias estimations observed in Table 2 is most probably inconsequential.

Insert Figure 5 about here

SEM variances and covariances. Figure 6 presents the distributional comparisons between the bootstrap bias estimator and the Monte Carlo bias estimator for four SEM variances and covariances [δ_{33} (CFA Model, N=200), ε_{44} (SEM Model 1, N=100), ϕ_{21} (SEM Model 1, N=500), and ε_{13} (SEM Model 2, N=200)]. Similar to Figure 5, the distributions of the bootstrap bias estimator and the Monte Carlo bias estimator are quite comparable in terms of their range and shape, indicating that the bootstrap bias estimation is adequate. Unlike the SEM fit indices discussed previously, for these (co)variances, there is no theoretical "ceiling" or "floor" which constrains the bias in certain direction for Monte Carlo bias estimation. The similar distributions of the bias estimators from the bootstrap and the Monte Carlo experiments suggest that the

difference between the bootstrap and the Monte Carlo bias estimations observed in Table 3 can be considered small and inconsequential. As pointed out before, it is somewhat difficult to judge the magnitude of bias for these variances and covariances because of their measurement scales.

Insert Figure 6 about here

SUMMARY AND CONCLUSIONS

This study examined the application of bootstrap method for bias estimation in the area of SEM. Three correctly specified SEM models were used under four different sample size conditions. Monte Carlo experiments were carried out to generate the criteria against which bootstrap bias estimation could be judged. For the Monte Carlo experiments, 200 replications were simulated within each cell condition, and the "true" bias of a statistic was obtained. For the bootstrap experiments, one parent bootstrap sample was generated under each cell condition, which served as the resampling space for bootstrapping within the cell. Two hundred bootstrap samples were obtained from this parent sample through sampling with replacement. Bootstrapped bias of a statistic was obtained and then compared with the bias of the statistic obtained from the Monte Carlo experiments. Both the amount and direction of bias, and the distributions of the bias estimator, from the bootstrap and Monte Carlo experiments were compared.

For SEM fit indices, bias estimation from the bootstrap and Monte Carlo experiments was quite comparable for most of them in terms of both the direction and magnitude. It is noted that, because true SEM models were simulated, bias was constrained in one direction in the Monte Carlo experiments, but less so in the bootstrap experiments. Because of this, bootstrap bias estimator has wider dispersion, and it was less skewed. It is hypothesized that if misspecified

SEM models were used in the study, the disparity between the bias estimator distributions from these two kinds of experiments would be reduced, because in this situation, the less-than-perfect population fit indices would function less as a constraint for the bias estimation in Monte Carlo experiments.

For the SEM loadings and coefficients, the difference between bias estimations from the bootstrap the Monte Carlo experiments was very small, although there did not appear to be a consistent pattern. The distributions of the bias estimators from the two experiments were quite similar both in terms of the shape and range, suggesting that the bootstrap bias estimation was adequate because its bias estimator was comparable to that from the Monte Carlo experiments.

For the SEM variances/covariances, the picture is similar to that about SEM loadings and coefficients. Because the measurement scales were quite different for these variances/covariances, caution was warranted in interpreting the magnitude of the estimated bias for these variances/covariances. The comparison of the distributions of the bias estimators from the two experiments indicated that bootstrap bias estimation could be considered adequate because of the similar distributions of the bias estimators from the two experiments.

To our knowledge, this is the first reported empirical study which provides direct empirical evidence with regard to the adequacy of bootstrap bias estimation in SEM. It is our hope that the findings from this study will make meaningful contribution to the SEM literature. Because the study involved three SEM models which served as internal replication mechanism, we are reasonably confident that the likelihood of chance discovery for the findings was small, and the findings should be generalizable. Future studies may extend the current findings by examining SEM models with some degree of misspecification, as discussed in this paper. In addition, data non-normality may be another dimension which should be considered in future investigation.

REFERENCES

Bentler, P. M. (1992). EQS structural equations program manual. Los Angeles, CA: BMDP Statistical Software.

Bentler, P. M. (1994). Forward. In B. M. Byrne, Structural Equation Modeling with EQS and EQS/Windows. Thousand Oaks, CA: SAGE.

Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. Annual Review of Psychology, 47, 563-592.

Bagozzi, R. P., Fornell, C., & Larcker, D. F. (1981). Canonical correlation analysis as a special case of structural relations model. Multivariate Behavioral Research, 16, 437-454.

Bollen, K. A. & Long, J. S. (1993). Introduction. In K. A. Bollen and J. S. Long (Eds.), Testing structural equation models (pp. 1-9). Newbury Park, California: SAGE Publications.

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. In C. C. Clogg (Ed.), Sociological methodology (pp. 115-140). Oxford: Basil Blackwell.

Bollen, K. A., & Stine, R. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In K. A. Bollen & J. S. Long, (Eds.), Testing structural equation models (pp. 111-135). Newbury Park, CA: Sage.

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. G. Jöreskog & H. Wold (Eds.), Systems under indirect observation: Causality, structure, prediction (Part I) (pp. 149-175). Amsterdam: North-Holland.

Byrne, B. M. (1994). Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming. Newbury, California: SAGE Publications, Inc.

Calsyn, J. R., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of

others: Cause or effect of academic achievement? Journal of Educational Psychology, 69, 136-145.

Camstra, A., & Boomsma, A. (1992). Cross-validation in regression and covariance structure analysis: An overview. Sociological Methods & Research, 21, 89-115.

Chatterjee, S. (1984). Variance estimation in factor analysis: An application of the bootstrap. British Journal of Mathematical and Statistical Psychology, 37, 252-262.

Dalgleish, L. I. (1994). Discriminant analysis: Statistical inference using the jackknife and bootstrap procedures. Psychological Bulletin, 116, 498-508.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7, 1-26.

Efron, B. (1985). Bootstrap confidence intervals for a class of parametric problems. Biometrika, 72, 45-48.

Efron, B. (1995). Computers, bootstraps, and statistics. Invited address to American Educational Research Association Annual Meeting, Division D--Measurement and Research Methodology. April, San Francisco, CA.

Fan, X. (1996). Structural equation modeling and canonical correlation analysis: What do they have in common? Structural Equation Modeling: A Multidisciplinary Journal, 4, 64-78.

Fan, X., & Jacoby, W. R. (1995). BOOTSREG: A SAS matrix language program for bootstrapping linear regression models. Educational and Psychological Measurement, 55, 764-768.

Fan, X., & Mathews, T. A. (1994, April). Using bootstrap procedures to assess the issue of predictive bias in college GPA prediction for ethnic groups. Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document

Reproduction Service #: ED 372 117).

Fan, X. & Wang, L. (1996). Comparability of jackknife and bootstrap results: An investigation for a case of canonical analysis. Journal of Experimental Education, 64, 173-189.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). Educational research: An introduction (6th ed.). White Plains, NY: Longman.

Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long, (Eds.), Testing structural equation models (pp. 40-65). Newbury Park, CA: SAGE Publications, Inc

Harris, D. J., & Kolen, M. J. (1989). Examining the stability of Angoff's delta item bias statistic using bootstrap. Educational and Psychological Measurement, 49, 81-87.

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? Psychological Bulletin, 112, 351-362.

Ichikawa, M., & Konishi, S. (1995). Application of the bootstrap methods in factor analysis. Psychometrika, 60, 77-93.

Jöreskog, K. G. & Sörbom, D. (1989). LISREL 7: A Guide to the Program and applications, 2nd Edition. Chicago, IL: SPSS Inc.

Kaiser, H.F., & Dickman, K. (1962). Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix. Psychometrika, 27, 179-182.

Kotz, S., & Johnson, N. L. (1992). Breakthroughs in statistics: Volumes 1 and 2. New York: Springer-Verlag.

Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1991). Approximating confidence intervals for factor loadings. Multivariate Behavioral Research, 26, 421-434.

Loehlin, J. C. (1992). Latent variable models: An introduction to factor, path, and structural

analysis. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. Psychological Bulletin, 111, 490-504.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. Psychological Bulletin, 103, 391-410.

Rasmussen, J. L. (1987). Estimating the correlation coefficients: Bootstrap and parametric approaches. Psychological Bulletin, 101, 136-139.

SAS Institute, Inc. (1990). SAS/STAT user's guide, version 6 (vol. 1, 4th ed.). Cary, NC: SAS Institute Inc.

Stine, R. A. (1989). An introduction to bootstrap methods: Examples and ideas. Sociological Methods and Research, 8, 243-291.

Tanaka, J. S. (1987). "How big is big?": Sample size and goodness-of-fit in structural equation models with latent variables. Child Development, 58, 134-146.

Thompson, B. (1988). Program FACSTRAP: A program that computes bootstrap estimates of factor structure. Educational and Psychological Measurement, 48, 681-686.

Thompson, B. (1993). The use of statistical significance tests in research: Bootstrap and other alternatives. The Journal of Experimental Education, 61, 361-377.

Wheaton, D. E., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), Sociological methodology 1977 (pp. 84-136). San Francisco: Jossey Bass.

Yung, Y., & Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), Advanced structural

equation modeling: Issues and techniques (pp. 195-226). Mahwah, New Jersey: Lawrence Erlbaum.

Table 1 Comparison of Bootstrap (Btrap) Bias Estimation and "True" Bias for SEM Fit Indices

MODEL[a]	N	SOURCE	χ^2	RMR[b]	GFI	AGFI	CFI	CENTRA	NNFI	NFI	RHO1	Δ2	PGFI	PARMS
CFA	100	Btrap	8.031	.033	-.022	-.079	-.025	-.037	-.064	-.024	-.059	-.025	-.009	-.009
		"True"	8.171	.147	-.026	-.091	-.008	-.014	-.020	-.022	-.055	-.008	-.010	-.009
	200	Btrap	6.779	.036	-.011	-.037	-.010	-.017	-.024	-.009	-.023	-.009	-.004	-.004
		"True"	8.508	.109	-.014	-.048	-.004	-.009	-.013	-.012	-.030	-.005	-.006	-.005
	500	Btrap	7.877	.022	-.005	-.018	-.004	-.008	-.011	-.004	-.011	-.004	-.002	-.002
		"True"	7.869	.064	-.005	-.018	-.001	-.005	-.007	-.004	-.011	-.003	-.002	-.002
	1000	Btrap	6.913	.015	-.002	-.008	-.002	-.003	-.005	-.002	-.005	-.002	-.001	-.001
		"True"	8.330	.047	-.003	-.010	-.001	-.004	-.006	-.002	-.006	-.002	-.001	-.001
SEM1	100	Btrap	47.808	1.713	-.054	-.090	-.056	-.193	-.079	-.049	-.069	-.055	-.039	-.035
		"True"	45.231	4.745	-.074	-.123	-.007	-.038	-.012	-.053	-.074	-.009	-.052	-.037
	200	Btrap	48.117	1.289	-.034	-.057	-.026	-.114	-.040	-.026	-.037	-.028	-.024	-.019
		"True"	48.767	3.429	-.038	-.064	-.003	-.028	-.009	-.026	-.037	-.006	-.027	-.019
	500	Btrap	49.339	.860	-.015	-.025	-.011	-.047	-.016	-.011	-.016	-.011	-.011	-.008
		"True"	48.473	2.119	-.016	-.026	-.001	-.025	-.008	-.011	-.015	-.006	-.011	-.008
	1000	Btrap	47.430	.569	-.008	-.013	-.005	-.023	-.007	-.005	-.007	-.005	-.005	-.004
		"True"	46.771	1.509	-.008	-.013	-.000	-.024	-.007	-.005	-.007	-.005	-.005	-.004
SEM2	100	Btrap	2.640	.071	-.007	-.037	-.012	-.012	-.043	-.010	-.036	-.011	-.002	-.003
		"True"	4.130	.178	-.013	-.070	-.005	-.003	-.008	-.017	-.063	-.002	-.004	-.004
	200	Btrap	4.095	.031	-.006	-.034	-.010	-.010	-.034	-.009	-.032	-.009	-.002	-.002
		"True"	3.994	.120	-.007	-.034	-.002	-.002	-.006	-.008	-.031	-.002	-.002	-.002
	500	Btrap	3.143	.025	-.002	-.011	-.002	-.003	-.010	-.003	-.010	-.003	-.001	-.001
		"True"	4.024	.076	-.003	-.014	-.001	-.002	-.007	-.003	-.013	-.002	-.001	-.001
	1000	Btrap	3.838	.017	-.001	-.007	-.002	-.002	-.006	-.002	-.006	-.002	-.000	-.000
		"True"	4.332	.056	-.001	-.008	-.001	-.002	-.007	-.002	-.007	-.002	-.000	-.001

a CFA: confirmatory factor analysis model; SEM1: structural equation model 1; SEM2: structural equation model 2.

b RMR: Root Mean-Square Residual; GFI: Goodness-of-fit index; AGFI: Adjusted goodness-of-fit index; CFI: comparative fit index; CENTRA: McDonald's centrality index; NNFI: Bentler-Bonnett's non-normed fit index; NFI: Bentler-Bonnett's normed fit index; RHO1: Bollen's Rho1; Δ2: Bollen's delta2 (incremental fit) index; PGFI: James et al.'s parsimonious goodness-of-fit index; PARMS: Mulaik et al.'s parsimonious fit index.

Table 2 Comparison of Bootstrap (Btrap) Bias Estimation with "True" Bias for SEM Loadings and Coefficients

Model	N	Source	λ_{X11}	λ_{X21}	λ_{X31}	λ_{X41}	λ_{X52}	λ_{X62}
CFA	100	Btrap	-.0040	.0002	.0191	-.0009	-.0021	-.0008
		"True"	.0004	-.0033	.0027	-.0011	-.0003	.0039
	200	Btrap	-.0030	.0084	.0062	.0060	-.0032	-.0003
		"True"	.0004	.0011	.0021	-.0028	-.0001	.0040
	500	Btrap	-.0002	-.0055	.0000	-.0013	-.0005	-.0026
		"True"	-.0001	-.0041	.0036	.0010	-.0004	-.0013
	1000	Btrap	-.0005	.0004	-.0027	-.0017	-.0010	-.0002
		"True"	.0001	.0008	-.0041	-.0021	-.0000	-.0006

Model	N	Source	λ_{X11}	λ_{X21}	λ_{X22}	λ_{X31}	λ_{X32}	λ_{X42}	λ_{Y11}	λ_{Y21}	λ_{Y22}	λ_{Y31}	λ_{Y32}	λ_{Y42}	γ_{11}	γ_{22}
SEM1	100	Btrap	.0037	-.0046	-.0082	.0000	.0015	.0097	-.0231	-.0287	-.0101	-.0308	-.0080	.0006	.0271	.0086
		"True"	-.0066	.0020	.0040	-.0006	.0082	.0012	.0003	-.0048	.0063	.0009	.0086	.0579	-.0002	-.0039
	200	Btrap	.0012	.0030	.0046	.0093	-.0018	.0083	-.0025	.0008	-.0005	.0047	-.0010	-.0045	.0015	.0044
		"True"	.0042	-.0032	.0015	.0000	-.0003	-.0059	.0152	.0052	-.0030	.0005	.0070	.0453	.0032	-.0018
	500	Btrap	.0001	-.0019	-.0017	-.0037	.0022	.0044	-.0014	-.0035	-.0012	-.0001	-.0021	.0003	.0021	.0019
		"True"	.0018	-.0019	.0004	-.0012	.0014	.0009	.0071	.0001	-.0002	.0029	-.0054	.0460	-.0013	.0031
	1000	Btrap	.0016	.0024	.0011	-.0001	.0001	.0010	-.0041	.0006	.0018	-.0017	.0011	-.0017	.0041	.0006
		"True"	.0026	-.0008	.0013	.0012	.0002	.0002	-.0007	.0020	-.0043	-.0003	-.0021	.0484	.0022	.0008

Model	N	Source	λ_{X11}	λ_{Y21}	λ_{Y42}	γ_1	γ_2	β
SEM2	100	Btrap	.1064	.0364	.0977	-.1007	-.1520	-.0240
		"True"	.0164	-.0141	.0055	-.0441	-.0274	-.0118
	200	Btrap	-.0199	-.0469	-.0304	-.0166	.0210	.0070
		"True"	.0109	.0002	-.0029	-.0023	.0016	.0013
	500	Btrap	.0031	.0035	.0011	-.0070	-.0013	-.0011
		"True"	-.0026	.0127	.0085	.0083	-.0005	-.0015
	1000	Btrap	.0012	.0006	.0019	-.0008	-.0070	-.0056
		"True"	.0022	-.0033	-.0032	-.0039	-.0027	.0005

Table 3 Comparison of Bootstrap (Btrap) Bias Estimation with "True" Bias for SEM Variances and Covariances

Model	N	Source	δ_{11}	δ_{22}	δ_{33}	δ_{44}	δ_{55}	δ_{66}	ϕ_{11}	ϕ_{22}	ϕ_{21}
CFA	100	Btrap	.048	-.051	-.120	-.044	.004	-.037	-.007	-.076	-.007
		"True"	-.067	-.019	-.016	-.004	-.032	-.038	.142	.022	.062
	200	Btrap	.018	-.047	-.024	-.049	-.033	-.012	-.097	-.018	-.068
		"True"	-.013	.029	-.020	.019	.001	-.011	-.006	-.098	-.052
	500	Btrap	-.011	-.014	.004	-.017	-.024	-.012	-.016	.008	.000
		"True"	.007	-.001	.006	-.015	-.008	-.002	.001	.022	.001
	1000	Btrap	-.017	-.017	-.017	-.012	-.005	-.001	-.002	-.006	-.001
		"True"	.003	.003	.004	-.001	-.006	.001	-.004	-.000	-.018

Model	N	Source	δ_{11}	δ_{22}	δ_{33}	δ_{44}	δ_{55}	δ_{66}	ε_{11}	ε_{22}	ε_{33}	ε_{44}	ε_{55}	ε_{66}	ϕ_{11}	ϕ_{22}	ϕ_{21}	ψ_{11}	ψ_{22}
SEM1	100	Btrap	-.407	-.528	-.289	-.517	-.761	-1.176	-.943	-1.236	-.652	-.127	-.659	-.996	-2.513	-1.418	-1.138	.950	.747
		"True"	-1.069	-1.205	-.694	.366	-.207	-.319	.444	-.471	-1.295	-.714	-.408	-.898	2.493	-.160	.056	.020	.029
	200	Btrap	-.792	-.287	-1.266	-.360	-.191	-1.205	.184	-.971	-.551	-.064	-.386	-.106	.217	-1.356	.366	.803	-.113
		"True"	-.453	-.489	-.543	-.381	-.095	.024	-.444	-.236	-.613	-.266	-.218	-.046	-.147	.993	-.706	-.789	.023
	500	Btrap	-.203	-.171	.009	-.074	.060	-.174	-.038	-.401	-.261	-.171	.009	-.350	-.382	-.690	-.273	.498	.077
		"True"	.391	-.230	-.455	-.012	-.024	-.146	-.266	-.006	.021	.129	-.204	.179	.468	.090	.391	-.489	.028
	1000	Btrap	-.119	-.155	-.322	-.027	-.152	-.070	-.038	-.018	-.087	.169	-.082	-.691	.065	-.233	.399	.065	.044
		"True"	-.336	-.040	-.258	.062	-.092	-.122	-.011	.034	-.092	-.343	-.340	.030	-.601	.494	.033	.071	.159

Model	N	Source	δ_{11}	δ_{22}	ε_{11}	ε_{22}	ε_{33}	ε_{44}	ϕ_{11}	ψ_{11}	ψ_{22}	ε_{13}	ε_{24}
SEM2	100	Btrap	.355	-.210	-.154	-.238	.315	-.622	-.413	-.355	-.760	.054	-.215
		"True"	.032	-.129	-.189	-.005	-.178	.017	.015	.072	-.083	-.010	-.032
	200	Btrap	-.520	.059	-.413	.341	-.418	.193	.506	.212	.138	-.271	.252
		"True"	-.033	-.039	-.104	.069	-.166	.024	.083	.025	-.003	-.139	.081
	500	Btrap	-.026	-.042	.030	-.054	-.028	-.029	.013	.002	-.039	.041	-.036
		"True"	-.071	.006	.069	-.074	.106	-.082	.092	-.029	-.019	.066	-.047
	1000	Btrap	-.053	-.004	-.045	-.029	-.028	.007	.050	-.006	-.048	-.004	-.006
		"True"	-.010	.003	-.054	.020	-.024	.003	-.005	.012	-.002	-.022	-.003

Table 4 Number of Samples with Improper Solutions

Sample N	Experiment	SEM Models		
		CFA	SEM1	SEM2
100	Bootstrap	15	17	41
	Monte Carlo	1	2	22
200	Bootstrap	0	0	29
	Monte Carlo	0	0	6
500	Bootstrap	0	0	3
	Monte Carlo	0	0	0
1000	Bootstrap	0	0	0
	Monte Carlo	0	0	0

Figure Captions

Figure 1. Confirmatory Factor Analysis Model and Model Specifications

Figure 2. Structural Equation Model 1 and Model Specifications

Figure 3. Structural Equation Model 2 and Model Specifications

Figure 4. Comparison of Bias Estimator Distributions for Four SEM Fit Indices (Black Bars: Bootstrap; Shaded Bars: Monte Carlo)

Figure 5. Comparison of Bias Estimator Distributions for Four SEM Loadings and Coefficients (Black Bars: Bootstrap; Shaded Bars: Monte Carlo)

Figure 6. Comparison of Bias Estimator Distributions for Four SEM Variances/Covariances (Black Bars: Bootstrap; Shaded Bars: Monte Carlo)

$$\Lambda_X = \begin{bmatrix} 1.00 & 0.00 \\ 0.85 & 0.00 \\ 0.80 & 0.00 \\ 0.70 & 0.00 \\ 0.00 & 1.00 \\ 0.00 & 0.75 \end{bmatrix} \qquad \theta_\delta \textit{ (Diagonal)} = \begin{bmatrix} 1.25 \\ 2.25 \\ 2.50 \\ 2.50 \\ 1.50 \\ 1.10 \end{bmatrix} \qquad \Phi = \begin{bmatrix} 4.00 & 5.00 \\ 5.00 & 9.00 \end{bmatrix}$$

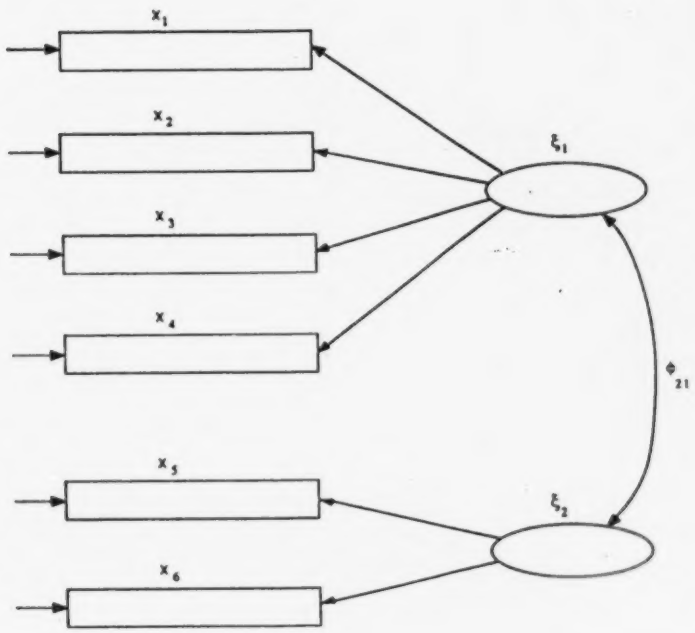


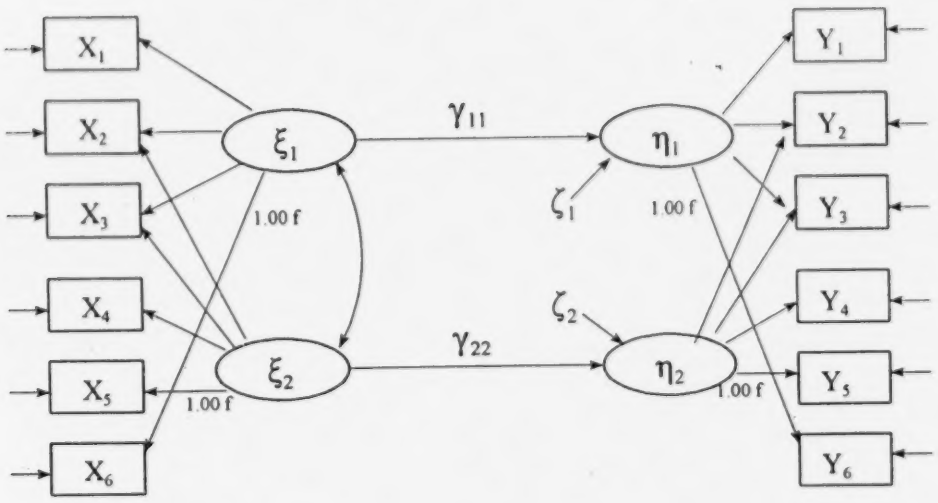
Figure 1

$$\Lambda_x = \begin{bmatrix} 0.95 & 0.00 \\ 0.20 & 0.50 \\ 0.25 & 0.85 \\ 0.00 & 1.20 \\ 0.00 & 1.00 \\ 1.00 & 0.00 \end{bmatrix} \qquad \Lambda_y = \begin{bmatrix} 0.90 & 0.00 \\ 0.25 & 0.65 \\ 0.35 & 0.70 \\ 0.00 & 0.95 \\ 0.00 & 1.00 \\ 1.00 & 0.00 \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} 0.90 & 0.00 \\ 0.00 & 0.80 \end{bmatrix} \qquad \Phi = \begin{bmatrix} 105.00 & 50.00 \\ 50.00 & 115.00 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix} \qquad \Psi = \begin{bmatrix} 18.00 & 0.00 \\ 0.00 & 6.00 \end{bmatrix}$$

$$\theta_\delta \text{ (Diagonal)} = \begin{bmatrix} 30.00 \\ 30.00 \\ 40.00 \\ 45.00 \\ 20.00 \\ 50.00 \end{bmatrix} \qquad \theta_\varepsilon \text{ (Diagonal)} = \begin{bmatrix} 25.00 \\ 40.00 \\ 50.00 \\ 40.00 \\ 20.00 \\ 70.00 \end{bmatrix}$$



1.00 f: Fixed to be 1

Figure 2

$$\Lambda_x = \begin{bmatrix} 1.00 \\ 0.50 \end{bmatrix} \qquad \Lambda_y = \begin{bmatrix} 1.00 & 0.00 \\ 0.95 & 0.00 \\ 0.00 & 1.00 \\ 0.00 & 0.90 \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} -0.60 \\ -0.25 \end{bmatrix} \qquad \Phi = [\, 7.00 \,]$$

$$B = \begin{bmatrix} 0.00 & 0.00 \\ 0.60 & 0.00 \end{bmatrix} \qquad \Psi = \begin{bmatrix} 5.00 & 0.00 \\ 0.00 & 4.00 \end{bmatrix}$$

$$\Theta_\delta = \begin{bmatrix} 3.00 & 0.00 \\ 0.00 & 2.50 \end{bmatrix} \qquad \Theta_\epsilon = \begin{bmatrix} 4.75 & & & \\ 0.00 & 2.50 & & \\ 1.60 & 0.00 & 4.50 & \\ 0.00 & 0.30 & 0.00 & 3.00 \end{bmatrix}$$

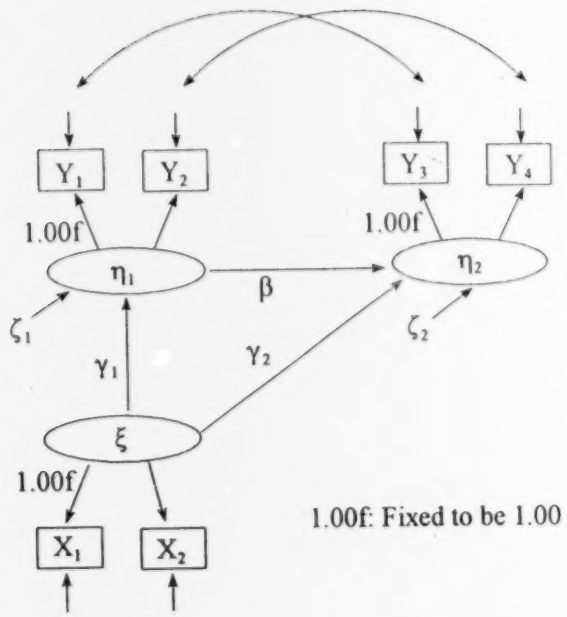
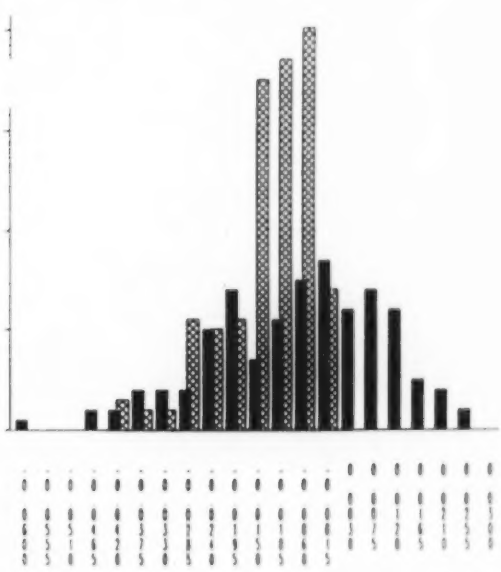
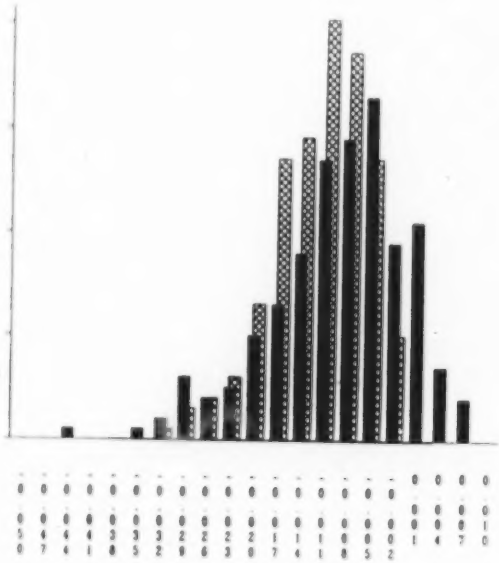


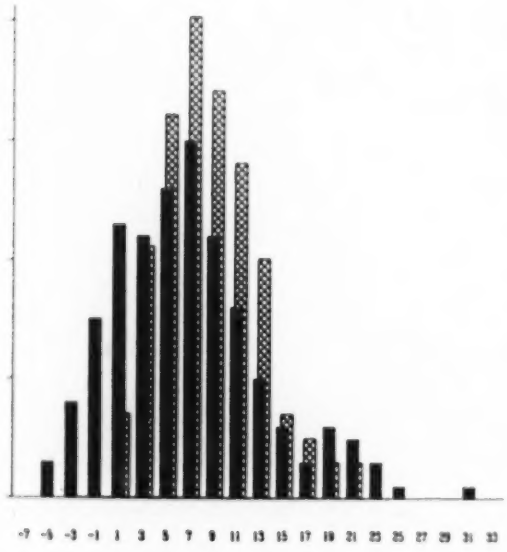
Figure 3



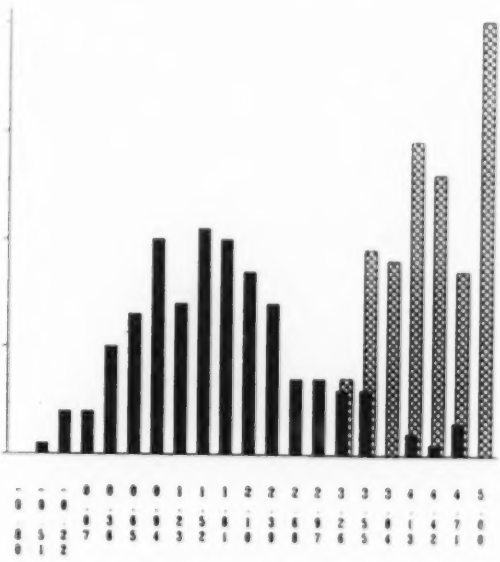
(a) GFI (SEM2, N=100)



(b) NFI (CFA, N=200)

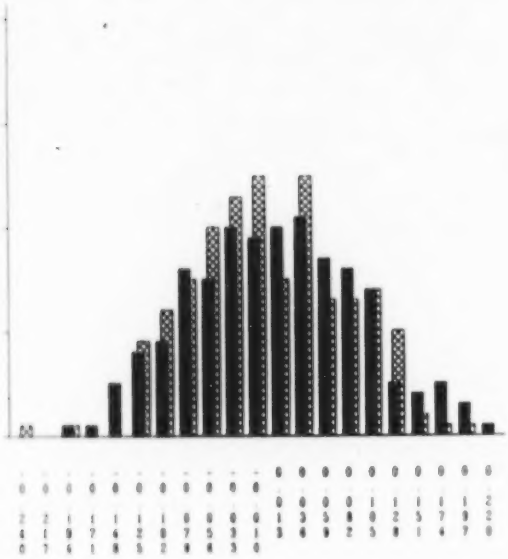


(c) χ^2 (CFA, N=200)

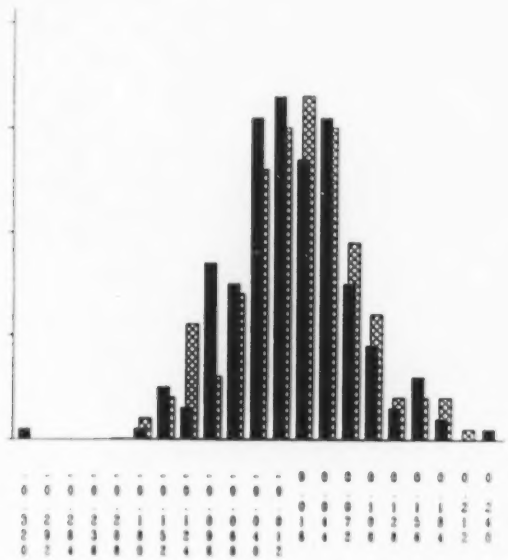


(d) RMR (SEM1, N=100)

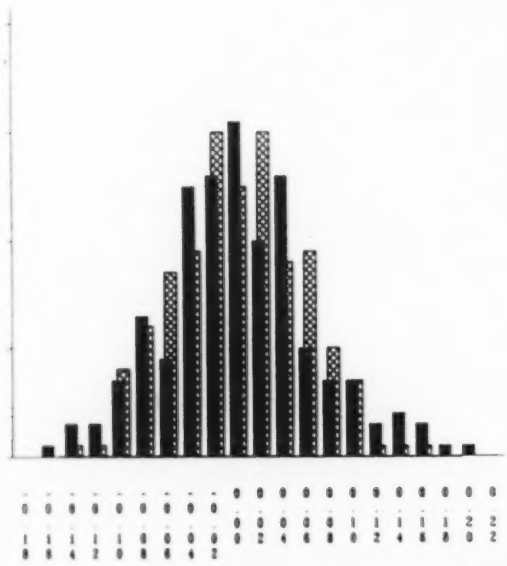
Figure 4



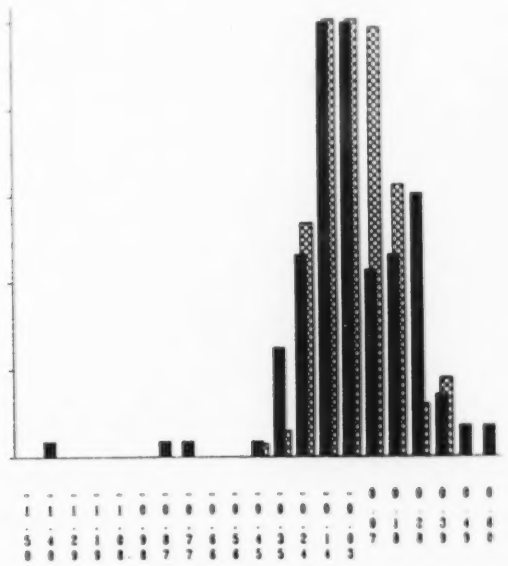
(a) λ_{X21} (CFA, N=200)



(b) λ_{Y32} (SEM1, N=200)

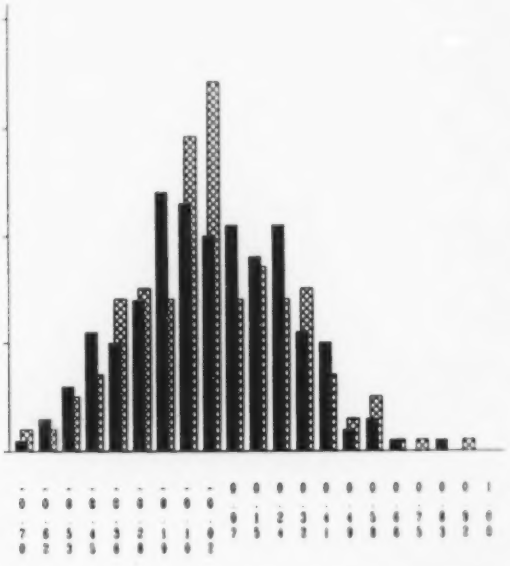


(c) γ_{11} (SEM1, N=500)

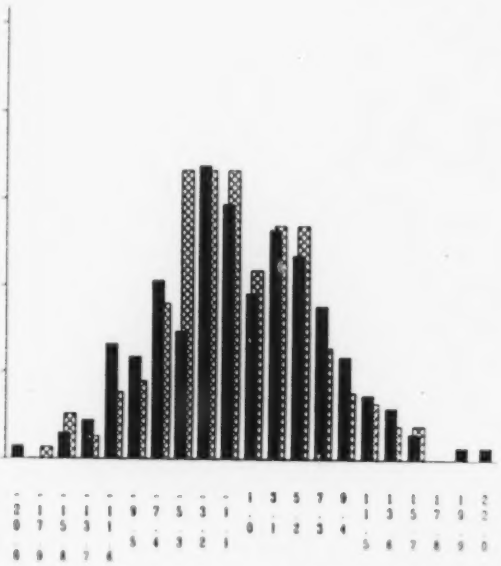


(d) β_{21} (SEM2, N=100)

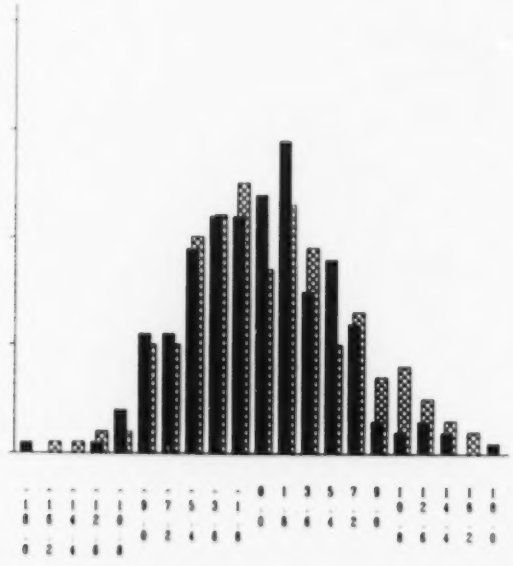
Figure 5



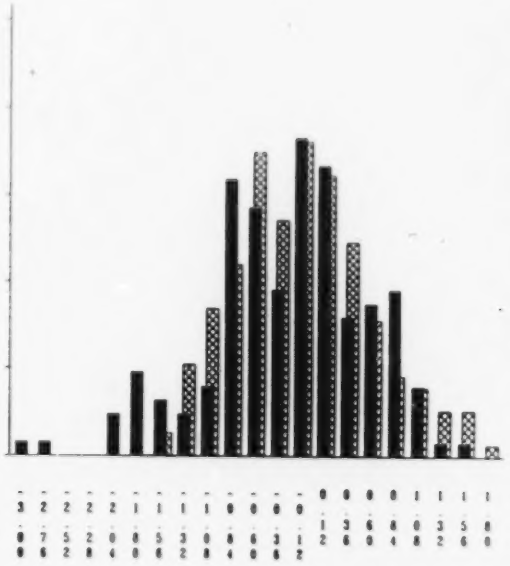
(a) δ_{33} (CFA, N=200)



(b) ε_{44} (SEM1, N=100)



(c) ϕ_{21} (SEM1, N=500)



(d) ε_{13} (SEM2, N=200)

Figure 6